

# Noncausality and Inflation Persistence

by

Markku Lanne\*

University of Helsinki

## Abstract

We use noncausal autoregressions to examine the persistence properties of quarterly U.S. consumer price inflation from 1970:1–2012:2. These nonlinear models capture the autocorrelation structure of the inflation series as accurately as their conventional causal counterparts, but they allow for persistence to depend on the size and sign of shocks to inflation as well as the inflation rate. Inflation persistence has decreased since the early 1980's, after which persistence is also greater following small and negative shocks than large and positive ones. At high levels of inflation, shocks are absorbed more slowly before the early 1980's and faster thereafter compared to low levels of inflation.

JEL Classification: C22, C51, E31

---

\*I thank the Academy of Finland and the OP-Pohjola Group Research Foundation for financial support. Part of this research was done while the author was visiting the German Institute for Economic Research in Berlin, whose hospitality is gratefully acknowledged.

## 1 Introduction

The persistence of inflation, i.e., the speed at which a shock to inflation is absorbed, contains important information for the conduct of monetary policy. In particular, the more persistent inflation is, the stronger the policy measures required to bring it back to the target level. As recently pointed out by Fuhrer (2010), a central challenge in inflation research is the mapping of observed or reduced-form persistence into the underlying economic structures that produce it. Part of this challenge comes from the fact that it is not even clear how reduced-form inflation persistence should be measured. A natural measure of persistence is the impulse response function implied by the inflation process, which shows how quickly the effect of a shock to inflation vanishes. Following the previous literature, Fuhrer (2010) emphasizes persistence measures based on autocorrelation, which are indeed reasonable if inflation dynamics are well described by a conventional (causal) autoregressive (AR) process, as is typically assumed. However, if this not the case, there is, in general, no one-for-one correspondence between the autocorrelation function and the (generalized) impulse response function.

In this paper, we consider the measurement of inflation persistence based on non-causal AR models previously shown to fit U.S. consumer price inflation well by Lanne and Saikkonen (2011), and Lanne, Luoma, and Luoto (2012). These models explicitly allow for dependence on (expected) future inflation, and, therefore, naturally facilitate gauging the relative importance of lagged inflation and future expectations in the determination of inflation. As the aforementioned authors point out, the typically observed strong autocorrelation does not necessarily translate into high persistence in the sense that inflation would strongly depend on past inflation, but it may rather be generated by agents' inflation expectations. With U.S. inflation, they indeed find expectations the prominent source of autocorrelation. However, they do not explicitly examine the persistence properties of inflation implied by the estimated noncausal AR models. As discussed in Section 2 below, while a causal AR process yields exactly the same autocorrelation function as its noncausal counterpart, the latter is nonlinear and, hence, allows for richer dynamics and, in particular, time-varying persistence.

Hence, despite its simplicity, the noncausal AR model facilitates examining changes in inflation persistence without having to pre-specify or test for any break dates, as has often been done in the previous literature on inflation persistence. In addition, it can yield information on the sources of potential changes in persistence because, unlike in the conventional causal AR model, the effect of a shock can depend on its sign, size and the level of inflation at the time when the shock occurs.

There are a number of theoretical economic models that imply asymmetries in inflation persistence. First, menu cost models may imply lower (higher) persistence following a large (small) or positive (negative) shock to inflation. As shown by Karadi and Reiff (2011), it is only after a large shock that a large number of firms have an incentive to change prices, leading to a quick response, while after small shocks prices are sticky as only few firms adjust their prices. Ball and Mankiw (1994), in turn, argue that after a positive inflation shock, firms have an incentive to increase their prices to keep their relative prices intact, while following a negative shock, they can save their menu costs by not changing their prices and letting the trend inflation reduce their relative prices. Second, if monetary policy reactions are asymmetric, persistence may depend on the sign of the shock to inflation. In particular, policy makers may be more concerned about negative than positive output gaps, indicating higher inflation persistence after a positive than negative inflationary shock (see, Gerlach (2000) and the references therein). Third, inflation persistence may also depend on the how far the inflation rate is from the target of the central bank. According to the opportunistic approach to disinflation put forth by Orphanides and Wilcox (2002) and Aksoy et al. (2006), the central bank reacts more aggressively when inflation is far from the target, indicating that at very high (or low) levels of inflation, shocks to inflation are absorbed more rapidly.

There seem to be very few previous empirical studies on inflation dynamics employing univariate nonlinear models, and, to the best of our knowledge, the persistence of U.S. inflation is systematically studied based on such a model only in the recent article by Nobay et al. (2010).<sup>1</sup> These authors fit an exponential smooth-transition

---

<sup>1</sup>There are, however, a number of attempts to examine the importance of nonlinearities for inflation persistence. First, Tsong and Lee (2011), and Tillmann and Wolters (2012) have applied

AR (ESTAR) model to quarterly U.S. PCE inflation from 1947:1 to 2004:4, and find that mean reversion is the faster the more the inflation rate deviates from equilibrium, and that large shocks are less persistent than small shocks.

While the ESTAR model can be motivated by results in recent monetary theory, it is also, to some extent, driven by these theories, which may be restrictive from the viewpoint of examining inflation persistence. In particular, the transition function of the ESTAR model is symmetric around equilibrium inflation, and hence, the persistence properties of positive and negative shocks of the same size cannot, by construction, be different. In the same vein, while the ESTAR model may imply different persistence properties dependent on the distance of the current inflation rate from the equilibrium, the differences are independent of whether the deviation is downward or upward. Neither of these restrictions is present in the noncausal AR model. Moreover, although linearity tests reject against the alternative of the ESTAR model, this may not be the true process because these tests have nontrivial power also against other nonlinear models. Specifically, Lof (2013) shows that STAR models may easily get mixed up with noncausal AR models considered in this paper. The noncausal AR model is also more general than the ESTAR model in at least two further ways. First, it has exactly the same autocorrelation function as its causal counterpart so that it captures the autocorrelation structure of inflation equally accurately. Second, as discussed by Lanne and Saikkonen (2013), it is capable of capturing GARCH effects prevalent in the inflation series, while Nobay et al. (2010) had to augment their 

---

the quantile-regression approach to causal AR models of inflation, which also facilitates studying whether positive and negative as well as large and small shocks have different effects. The former authors conclude that large negative shocks are absorbed more quickly than positive shocks in twelve OECD countries, while the latter authors find a break in the persistence of the postwar U.S. inflation rates such that since the 1980's inflation has become less persistent at all quantiles. Second, structural breaks have been introduced into causal AR models for inflation. Levin and Piger (2003), and Cecchetti and Debelle (2006) have shown that if breaks in the intercept of a causal AR model are neglected, inflation may spuriously appear too persistent. Third, Kumar and Okimoto (2007) and Hassler and Meller (2011) estimate fractionally integrated AR models for U.S. consumer price inflation. The former authors find much lower persistence after 1982, while according to Hassler and Meller (2011) the break in the early 1980's is not statistically significant.

ESTAR model with a separate GARCH component to get an adequate description of inflation dynamics.

In Section 4, we estimate a noncausal AR model for quarterly U.S. consumer price inflation from 1970:1 to 2012:2, and compute generalized impulse response functions based on it. In addition to the entire sample period, we report results separately for two subsample periods, 1970:1–1982:4 and 1983:1–2012:2. In accordance with much of the previous literature (see Fuhrer (2010) and the references therein), we find a clear reduction in inflation persistence since the early 1980’s. Moreover, large and negative shocks appear more persistent than small and positive ones in the post-1982 period, while no such clear nonlinear effects prevail in the pre-1983 period. Before 1983, shocks are absorbed more slowly when the inflation rate is high, whereas the case seems to be the opposite in the more recent subsample period, potentially suggesting that monetary policy aimed at keeping inflation under control has become more aggressive.

The plan of the rest of the paper is as follows. In Section 2, we describe the noncausal AR model and discuss model selection. Section 3 deals with the computation of the generalized impulse response function. Empirical results are reported in Section 4. Finally, Section 5 concludes.

## 2 Noncausal Autoregression

The starting point of our analysis is the noncausal AR model of Lanne and Saikkonen (2011) that can be described as follows.<sup>2</sup> Consider a stochastic process  $y_t$  ( $t = 0, \pm 1, \pm 2, \dots$ ) generated by

$$\phi(B)\varphi(B^{-1})y_t = \epsilon_t, \quad (1)$$

where  $\phi(B) = 1 - \phi_1 B - \dots - \phi_r B^r$ ,  $\varphi(B^{-1}) = 1 - \varphi_1 B^{-1} - \dots - \varphi_s B^{-s}$ ,  $\epsilon_t$  is a sequence of independent, identically distributed (continuous) random variables with

---

<sup>2</sup>An alternative formulation is proposed by Breidt et al. (1991). However, as Lanne and Saikkonen (2011) point out, their model has the advantages that it is straightforward to test for the specified number of leads and lags and inference on the autoregressive parameters is asymptotically independent of inference on the parameters of the error distribution.

mean zero and variance  $\sigma^2$  or, briefly,  $\epsilon_t \sim i.i.d. (0, \sigma^2)$ , and  $B$  is the usual backward shift operator, i.e.,  $B^k y_t = y_{t-k}$  ( $k = 0, \pm 1, \dots$ ). The polynomials  $\varphi(z)$  and  $\phi(z)$  are assumed to have their zeros outside the unit circle so that

$$\phi(z) \neq 0 \quad \text{for } |z| \leq 1 \quad \text{and} \quad \varphi(z) \neq 0 \quad \text{for } |z| \leq 1. \quad (2)$$

We use the abbreviation  $\text{AR}(r, s)$  for the model defined by (1). If  $\varphi_1 = \dots = \varphi_s = 0$ , model (1) reduces to the conventional causal  $\text{AR}(r, 0)$  model with  $y_t$  depending on its past but not future values. The more interesting cases arise, when this restriction does not hold. If  $\phi_1 = \dots = \phi_r = 0$ , we have the purely noncausal  $\text{AR}(0, s)$  model with dependence on future values only. In the mixed  $\text{AR}(r, s)$  case where neither restriction holds,  $y_t$  depends on its past as well as future values.

To gain further insight, let us write model (1) as

$$y_t = \phi_1 y_{t-1} + \dots + \phi_r y_{t-r} + \sum_{j=0}^{\infty} \beta_j E_t(\epsilon_{t+j}), \quad (3)$$

where the conditional expectation is taken with respect to past and current values of the process (see Eq. (6) in Lanne and Saikkonen (2011)). Unlike in the causal AR model, current and future errors are predictable by the history of the process, and, thus, the last term in (3), in general, differs from zero. Hence, if  $y_t$  is noncausal, its dynamics are at least in part driven by expected future errors.

A well-known feature of noncausal autoregressions is that a non-Gaussian error term is required to achieve identification. Therefore, we assume that  $\epsilon_t$  in (1) is non-Gaussian and its distribution has a (Lebesgue) density  $f_\sigma(x; \boldsymbol{\omega}) = \sigma^{-1} f(\sigma^{-1}x; \boldsymbol{\omega})$  which depends on the parameter vector  $\boldsymbol{\omega}$  ( $d \times 1$ ) in addition to the scale parameter  $\sigma$  already introduced. The function  $f(x; \boldsymbol{\omega})$  is assumed to satisfy the regularity conditions stated in Andrews et al. (2006), and Lanne and Saikkonen (2011). These conditions imply that  $f(x; \boldsymbol{\omega})$  is twice continuously differentiable with respect to  $(x, \boldsymbol{\omega})$ , non-Gaussian, and positive for all  $x \in \mathbb{R}$  and all permissible values of  $\boldsymbol{\omega}$ . Following, Lanne and Saikkonen (2011), and Lanne, Luoma, and Luoto (2012), we use Student's  $t$  distribution as the error distribution for the U.S. inflation in Section 4. There has recently been increasing interest in non-Gaussian distributions in the

context of macroeconomic time series; for instance, Chib and Ramamurthy (2011) and Cúrdia et al. (2012) have considered macroeconomic models with  $t$ -distributed errors.

Lanne and Saikkonen (2011) showed how model (1) can be consistently estimated by the method of maximum likelihood (ML). They also showed that the (local) ML estimator is asymptotically normally distributed, and a consistent estimator of the limiting covariance matrix is obtained in the usual way from the standardized Hessian of the log-likelihood function. Thus, standard errors of estimators and conventional Wald test statistics with an asymptotic  $\chi^2$ -distribution under the null hypothesis can be computed as usual.

Once noncausality is allowed for, selecting the correct specification among the different  $AR(r,s)$  models adequately capturing the autocorrelation in the series of interest becomes an issue. Breidt et al. (1991) suggested a model selection procedure based on maximizing the likelihood function, where all purely noncausal, causal and mixed models of a given order  $p$  are estimated, and the model yielding the greatest value of the likelihood function is selected. The order  $p$  is first selected by finding a causal Gaussian AR model exhibiting serially uncorrelated errors. The simulation results of Lanne and Saikkonen (2011) indicate that this procedure works very well in small samples. Nevertheless, they recommend augmenting it with diagnostic checks concerning remaining autocorrelation in the residuals and their squares as well as the fit of the proposed error distribution. As pointed out in the Introduction, the correctly specified noncausal AR model is capable of capturing conditional heteroskedasticity. Hence, if the true model is noncausal, the residuals of a fitted causal or misspecified noncausal AR model tend to appear heteroskedastic, indicating that autocorrelation of the squared residuals is likely to be useful in detecting misspecification.<sup>3</sup>

---

<sup>3</sup>In Section 4, we use conventional portmanteau tests and quantile-quantile (Q-Q) plots as diagnostic checks. It should, however, be noticed that, when the orders of the model are misspecified, the errors are dependent, and, therefore, the portmanteau tests are not exactly valid as they do not take the estimation error correctly into account. Nevertheless, p-values of these tests can be seen as convenient summary measures.

### 3 Generalized Impulse Response Function

As pointed out in the Introduction, the impulse response function (IRF) of a process conveniently summarizes its persistence properties by providing a measure of the effect of a shock  $v_t$  occurring at time  $t$  on the time series  $y_t$  after  $n$  periods. For a linear AR model, its computation is straightforward, but with nonlinear noncausal AR models, a number of complications arise. In particular, the impulse responses can depend on the size and sign of the shock  $v_t$  as well as the information used to forecast future values of  $y_t$ . For a noncausal  $\text{AR}(r, s)$  model, this information, typically called ‘history’ and denoted by  $\omega_{t-1}$ , consists of the  $r+s$  past values of the process,  $y_{t-1}, \dots, y_{t-(r+s)}$ . The generalized impulse response function (GIRF) introduced by Koop et al. (1996) offers a solution to these complications. For a specific shock  $\delta_t$  and history  $\omega_{t-1}$ , it is defined as

$$\text{GIRF}(n, \delta_t, \omega_{t-1}) = E(y_{t+n} | v_t = \delta_t, \omega_{t-1}) - E(y_{t+n} | v_t = 0, \omega_{t-1}), \quad (4)$$

for  $n = 1, 2, \dots$ . This yields the generalized impulse responses of a shock of size  $\delta_t$  conditional on a given history  $\omega_{t-1}$ . The GIRF of the causal  $\text{AR}(r, 0)$  model coincides with its conventional IRF. For the noncausal  $\text{AR}(r, s)$  model with  $s > 0$ , the conditional expectations cannot be computed in closed form or using the simulation method suggested by Koop et al. (1996) and commonly employed for computing forecasts from nonlinear models. However, relevant simulation-based forecasting methods are provided by Lanne, Luoto, and Saikkonen (2012), and Lanne, Luoma, and Luoto (2012), the former of which is employed in Section 3.

As discussed by Koop et al. (1996), the GIRF can be computed with several alternative assumptions, depending on the interests of the researcher. For instance, the shock  $\delta_t$  can be drawn from some distribution, in which case, for a fixed history  $\omega_{t-1}$ , GIRF becomes a random variable in terms of the shocks. Alternatively, if the shock is fixed and the histories are drawn from, say, the set of observed histories, we get a random variable in terms of the history. In Section 4, we will compute the typical GIRF for each quarter by fixing the history and the shock to the  $r+s$  preceding observed inflation rates and the one-step forecast error at each data point,

respectively. The latter measures the part of  $y_t$  that is unpredictable by past inflation, and, hence, it can be interpreted as the most typical shock at each time point. The residuals of the estimated model are often used to measure shocks (see, e.g., Pesaran and Potter (1997)), but because the residuals of the noncausal AR model depend on future values of  $y_t$ , they are not unpredictable, and thus not useful proxies of the shocks.

The simplest case where the properties of the GIRF of noncausal AR models can be studied, is the AR(0,1) model where the relevant history only depends on one past value. Our experimentation with this model indicates that its GIRF indeed can strongly depend on both its history and the size of the shock, while the effect of the sign of the shock is minor. Comparison with the IRF of the corresponding conventional causal AR(1,0) model further reveals that depending on the history and the shock, the AR(0,1) model can exhibit much less or somewhat more persistence than its causal counterpart.

#### 4 Empirical Results

We model the (demeaned) annualized quarterly inflation rate computed from the seasonally adjusted U.S. consumer price index (CPI) for all urban consumers from 1970:1 to 2012:2. The source of the data is the FRED database of the Federal Reserve Bank of St. Louis. We choose this particular inflation series because noncausal AR models have previously been shown to fit well the same series spanning a somewhat shorter period by Lanne and Saikkonen (2011), and Lanne, Luoma, and Luoto (2012). Specifically, both papers present strong evidence in favor of noncausality, and, in addition, the latter authors find the noncausal AR models clearly superior to their causal counterparts in forecasting. Following the aforementioned papers, the sample period only starts in 1970 because it does not seem possible to specify an adequate AR model for the entire series that is available from 1947.

The inflation series and its autocorrelation function are plotted in Figure 1. The periods of increased volatility in the 1970's and early 1980's as well as the subsequent period of Great Moderation are clearly discernible. The outlier of the fourth quarter

of 2008 is noteworthy, and we will later check the robustness of the results with respect to it. Positive autocorrelation is relatively strong even at high lags, which has typically been interpreted as a sign of high inflation persistence in the previous literature.

#### 4.1 Model Selection

The first step in the analysis is model selection considered in this subsection. In other words, we first find a (potentially noncausal) AR model that adequately captures the dynamics of the inflation series. Persistence is then studied based on the selected model in Subsection 4.2.

The estimation results of a number of AR models are presented in Table 1. Following Lanne and Saikkonen (2011), we start out by first specifying a causal Gaussian AR model that adequately captures the autocorrelation in the inflation series, and check its residuals for normality. The third-order model (AR(3,0)- $N$ ) seems adequate in terms of both no remaining autocorrelation and conditional heteroskedasticity. The quantile-quantile (Q-Q) plot of the residuals in the upper panel of Figure 2 indicates clear inadequacy of the Gaussian model. In particular, there are great discrepancies at the tails, suggesting the need for a fat-tailed error distribution, such as Student's  $t$  distribution with a small value of the degree-of-freedom parameter.

The four rightmost columns of Table 1 present the results of the third-order AR models with a Student- $t$  error distribution. The estimates of the degree-of-freedom parameter  $\lambda$  of all models are quite small in accordance with the fat tails of the error distribution suggested by the Q-Q-plot of the causal Gaussian AR model. The preferred model is the purely noncausal specification (AR(0,3)- $t$ ) that maximizes the log-likelihood function by a sizable margin compared to the other models. All parameters are estimated very accurately. The AR(0,3)- $t$  model seems adequate in that the errors exhibit no autocorrelation or conditional heteroskedasticity, and according to the Q-Q-plot in the lower panel of Figure 2, the Student- $t$  distribution indeed seems to provide quite a good fit, also at the tails. We also checked the adequacy of the AR(0,3)- $t$  specification by testing the significance of the additional lag and lead in

AR(1,3)- $t$  and AR(0,4)- $t$  models, respectively. Both turned out insignificant at the 5% level (the p-values are 0.067 and 0.539 for the lag and lead, respectively). Hence, U.S. consumer price inflation seems to be driven by expectations of future errors (see Eq. (3)), as also found by Lanne and Saikkonen (2011), and Lanne, Luoma, and Luoto (2012).

Because the potential outlier of 2008:4 may affect model selection and the diagnostic tests, in order to check for robustness, we also considered a series adjusted by replacing this observation by the average of the adjacent observations. The AR(3,0)- $N$  model adequately captures the autocorrelation of this series as well, but the McLeod-Li test rejects at the 10% significance level. The vast reduction in the p-value is not surprising as it is well known that the presence of an outlier can lead to overrejection and low power of ARCH tests (see, e.g., van Dijk et al. (1999)). Among the third-order AR specifications, the AR(0,3)- $t$  model is selected, with even a wider margin. Moreover, the p-values of the diagnostic tests for the selected model are higher, attesting to the effect of the aberrant observation on these tests. Nevertheless, the conclusions drawn above seem robust with respect to the potential outlier, and we may proceed with the selected AR(0,3)- $t$  model.

Before proceeding to impulse response analysis, we finally check whether the inflation series is a unit root process. In view of the evidence of noncausality, it would be consistent to base the test on the selected noncausal AR(0,3)- $t$  model. However, to the best of our knowledge, a unit root test allowing for noncausality is not yet available, and, therefore, we use the conventional augmented Dickey-Fuller test based on a regression of the difference of  $y_t$  on  $y_{t-1}$ , two lagged differences and an intercept, corresponding to the third-order AR model deemed adequate. The value of the test statistic is  $-2.91$  with a p-value of 0.047, indicating the rejection of the unit root hypothesis at the 5% level of significance.<sup>4</sup>

---

<sup>4</sup>Pentti Saikkonen and Rickard Sandberg are currently working on a unit root test for noncausal AR processes, and according to a private communication, their preliminary results show that a great power loss may result if the test is incorrectly based on the causal model. Hence, a rejection at even lower significance levels would be expected if a test based on the correct model were used.

## 4.2 Impulse Response Analysis

In this subsection, we examine inflation persistence by means of the generalized impulse response function. As discussed in Section 3, the relevant shock whose dynamic effects we examine, is the one-step ahead forecast error that is unpredictable by past inflation. Figure 3 contains a box-and-whisker plot describing the densities of these forecast errors. To make the figure legible, the density of the very large forecast error (median  $-20.50$ ) pertaining to the potential outlier of 2008:4 is not shown. It is clear that the density of the typical shock varies a lot from quarter to quarter so that drawing them from a fixed distribution or setting the shock at a fixed value would not produce GIRFs with a meaningful interpretation. For instance, unity is not included in the 90% range in all quarters. Moreover, the upper panel of Figure 1 shows that the relevant histories also change considerably over time, and visual inspection of the two figures together suggests that shocks may be history-dependent. Specifically, small shocks with low variance seem to be typical in the period of the Great Moderation, when inflation was low and varied little. In the volatile period of the early 1980's, on the other hand, large shocks with greater variance seem more common. Hence, computing the typical GIRF for each quarter by fixing the history and the shock should be the most informative way to proceed.

Visual inspection of the typical impulse response functions shows that they indeed vary a lot in shape from quarter to quarter. The decline to zero is never monotonous, and the implied persistence seems to vary considerably. When reporting the results below, we condition on the size and sign of the shock as well as the level of inflation at the time of the shock by computing averages over the relevant GIRFs.

To summarize the persistence properties, we report the average  $\pi$ -life (or absorption time) of a shock, conditional on the absolute size of the shock and the level of inflation at the time of the shock. This measure of persistence suggested by van Dijk et al. (2007) is related to the half-life measure commonly used in the previous empirical literature, but it is more general. In particular, the half-life defined as the minimum time it takes for the impulse response function to reach a value less than half of the initial impact is, in general, meaningful only when the impulse response

function decays monotonically, which is not the case here.

In short, the  $\pi$ -life  $N(\pi, \delta_t, \omega_{t-1})$  related to a generalized impulse response function  $GIRF(n, \delta_t, \omega_{t-1})$ ,  $n = 1, 2, \dots$ , gives the minimum value of  $n$  beyond which the difference between the impulse responses at all greater values of  $n$  and the ultimate response ( $GIRF^\infty(\delta_t, \omega_{t-1})$ ) is less than or equal to a fraction  $\pi$  of the difference between the initial impact and the ultimate response. Formally it can be written as

$$N(\pi, \delta_t, \omega_{t-1}) = \sum_{m=0}^{\infty} \left[ 1 - \prod_{n=m}^{\infty} I(\pi, n, \delta_t, \omega_{t-1}) \right], \quad (5)$$

where the indicator function

$$I(\pi, n, \delta_t, \omega_{t-1}) \equiv I[|GIRF(n, \delta_t, \omega_{t-1}) - GIRF^\infty(\delta_t, \omega_{t-1})| \leq \pi |\delta_t - GIRF^\infty(\delta_t, \omega_{t-1})|]$$

for  $0 \leq \pi \leq 1$  (for further details, see van Dijk et al. (2007)). In practice, a finite maximum value of periods to the ultimate response must be chosen to approximate  $GIRF^\infty(\delta_t, \omega_{t-1})$  and to truncate the summation in (5). Following the advice of van Dijk et al. (2007), we choose a fairly large truncation point of 40, where the values of the GIRF are already very close to zero. Because the  $\pi$ -life only takes integer values, it may not always be able to discriminate between processes with different persistence properties. To remedy this, van Dijk et al. (2007) recommend computing it for several values of  $\pi$ .

The average  $\pi$ -lives for various values of  $\pi$  and conditioned on the absolute shock size are reported in Table 2. As is common in the empirical literature on U.S. inflation, we also report the results for two subsample periods, 1970:1–1982:4, and 1983:1–2012:2, the latter of which is characterized by a large reduction in the volatility of inflation and a new policy regime. The starting date of the new regime is not unambiguous, but, for comparability, we choose to use the same division as Nobay et al. (2010), whose study comes closest to ours in the previous literature. The leftmost column (marked with heading A) corresponding to each subsample period gives the average absorption time over all shocks as a function of  $\pi$ . The absorption times naturally increase as  $\pi$  decreases, but the general conclusion, irrespective of the value

of  $\pi$ , is that persistence is much lower in the latter subsample period. The same conclusion was also drawn by Nobay et al. (2010) for the PCE inflation and Benati (2008), inter alia, for consumer price inflation. The absorption times implied by the causal AR(3,0)- $t$  and AR(3,0)- $N$  models (not shown) are similar in magnitude, and actually indicate somewhat lower persistence when  $\pi = 0.5$ . However, as they are independent of the size of the shock, they are unable to reveal the nonlinearities in persistence that we next turn to.

In order to gauge the effect of the shock size, the quarters of each sample period are divided into three groups of approximately equal magnitude by the absolute size of the shock. In the latter subsample period, the shocks are considerably smaller on average, which, of course, shows up in the cutoff points between the groups. They are 1.71 and 0.69 between the small and medium-sized shocks, and 2.53 and 1.56 between the medium-sized and large shocks in the two periods, respectively. In each group, the average  $\pi$ -life is always clearly greater in the first subsample period, reconfirming the findings based on all shocks taken together. Moreover, in the latter subsample period, the persistence always decreases as the size of the shock increases. In the first subsample period there is a similar tendency although it is not equally clear. This is in line with the implications of the menu cost models discussed in the Introduction as well as the results of Nobay et al. (2010) although they did not find much difference between the subsample periods. A potential explanation to this slight discrepancy might be that they considered shocks of the same fixed size in both periods, while our GIRFs are based on shocks typical in each quarter.

Table 3 shows how the average absorption times depend on the level of inflation at the time of the shock. The quarters of each sample period are divided into three groups of approximately equal magnitude by the inflation rate. In the entire sample period, the dependence on the initial inflation is rather weak and seems to be masked by the fact that the differences between the subsample periods are rather large and the reactions go in different directions. The differences between the low and medium inflation quarters are minor within and between the subsample periods. However, in the quarters with high inflation, the absorption times are much longer (two to three times, depending on  $\pi$ ) in the first subsample period compared to the post-1982

period. The respective cutoff points between medium and high inflation are 8.49 and 3.55, suggesting that in the first subsample period, a typical shock took a much longer time to have an effect despite the higher level of inflation, while in the post-1982 period, a shock was absorbed quickly when inflation was relatively high. These results lend some support to the opportunistic approach to disinflation discussed in the Introduction. In particular, in the latter subsample period, a potential explanation to the more rapid absorption of shocks in quarters of high inflation may be more aggressive monetary policy aimed at keeping inflation under control. Moreover, the fact that it is only at high levels of inflation that there is less persistence, suggests that the central bank is more concerned about inflation much above than below the target. Our findings are also not entirely in conflict with those of Nobay et al. (2010), who concluded that the speed of adjustment is the faster the further away inflation is from its mean in the entire sample, and such mean reversion is more rapid in the latter subsample period. However, the transition function of the ESTAR model is by construction symmetric, implying a faster absorption of shocks both at low and high inflation levels.

In addition to the examination of the dependence of the absorption times of shocks on their size and the current inflation rate, it is interesting to check whether positive and negative shocks have different effects. To that end, we employ the asymmetry measure of van Dijk et al. (2007), defined for a specific shock  $\delta_t$  and history  $\omega_{t-1}$  as the difference between the  $\pi$ -lives of  $\delta_t$  and  $-\delta_t$ ,

$$ASYN(\pi, \delta_t, \omega_{t-1}) = N(\pi, \delta_t, \omega_{t-1}) - N(\pi, -\delta_t, \omega_{t-1}). \quad (6)$$

If the absorption of the shock  $\delta_t$  is symmetric, this measure equals zero for all values of  $\pi$ . Otherwise, assuming  $\delta_t$  is positive (negative), a positive (negative) value indicates that a positive (negative) shock takes a longer time to be absorbed than a negative (positive) shock. It is also possible to assess statistically whether asymmetry is significant by testing whether the average asymmetry measure over (a subset of) shocks and histories equals zero. Unfortunately, this is hampered by the fact that the realizations of  $ASYN(\pi, \delta_t, \omega_{t-1})$  are not independent across histories, as pointed out by van Dijk et al. (2007). However, they suggest a conservative estimate of the

standard error obtained by dividing the standard deviation of the asymmetry measure by the square root of the number of combinations of shocks  $\delta_t$  and histories  $\omega_{t-1}$  over which it is computed. In our case, each history corresponds to one typical shock, so the latter equals the number of shocks included in the subset of interest.

In Table 4 are presented the average asymmetry measures for negative and positive shocks. We compute the  $\pi$ -life pertaining to the opposite of each shock, and obtain the asymmetry measure for each quarter using (6). The asymmetry measures are then averaged separately over positive and negative shocks in each sample period. Hence, positive (negative) entries in the table tell how much shorter (longer) the average  $\pi$ -life would be if the shock were the opposite of the typical shock in each quarter. The positive and negative shocks are on average approximately equal in absolute value in each sample period. Following van Dijk et al. (2007), an average asymmetry measure exceeding twice its standard error is considered significant. Asymmetry is greatest for the positive shocks in the latter subsample period, which is also reflected in the results for the entire sample period, and only these asymmetry measures are deemed significant. These negative values indicate that in the quarters when the typical shock is positive, a negative shock of the same magnitude would have been more persistent. This contrasts the view discussed in the Introduction that the central bank is more concerned about negative than positive output gaps, making positive inflationary shocks more persistence than negative ones. On the contrary, one potential explanation to our finding is that the central bank now reacts more aggressively to positive than negative deviations from the inflation target. This is also consistent with the finding in Table 3 that in the latter subsample period, shocks are absorbed more rapidly when the inflation rate is high than when it is low.

## 5 Conclusion

In this paper, we study the persistence properties of quarterly U.S. consumer price inflation by means of noncausal AR models, shown to fit and forecast inflation series well in the previous literature. While these models capture the autocorrelation structure as accurately as their conventional causal counterparts, they are nonlinear,

and, hence, can allow for different persistence properties of shocks of different size and sign. Moreover, the absorption of shocks can depend on the inflation rate at the time of the shock. A number of theoretical models imply these kinds of nonlinearities, but the noncausal AR model is general in that it does not embody the theoretical restrictions at the outset, and thus, facilitates more flexible analysis than, say, the ESTAR model previously used to examine similar effects.

We consider quarterly U.S. consumer price inflation from 1970:1 to 2012:2. In addition to the entire sample period, we report results separately for two subsample periods, 1970:1–1982:4 and 1983:1–2012:2. In accordance with much of the previous literature, we find a clear reduction in inflation persistence since the early 1980’s. Moreover, large and negative shocks seem to be more persistent than small and positive ones only after 1982. The dependence of inflation persistence on the inflation rate also seems to have changed in the early 1980’s. Before 1983, high persistence is linked to a high inflation rate, while the situation appears to be the opposite after 1982. This suggests that monetary policy, aimed at keeping inflation under control has become more aggressive.

This paper is, to the best of our knowledge, the first attempt to study the persistence properties of inflation by means of noncausal AR models. Compared to the previous literature on generalized impulse response analysis, a central challenge with these models is the definition of the shock. Our solution is to measure the typical shock for each quarter by the one-period forecast error, but it might be interesting to examine the robustness of this choice. We also assume that the parameters of the model are constant over time. Despite this assumption, with the noncausal AR model, the persistence properties are still allowed to vary over time. However, the model could be extended to allow for breaks or time-varying parameters. Finally, while we have only considered CPI inflation, it might be interesting to examine whether similar conclusions hold for other U.S. and international inflation series, including PCE or GDP deflator inflation. We leave these issues for future research.

## References

- Aksoy, Y., A. Orphanides, D. Small, V. Wieland, and D. Wilcox (2006). A quantitative exploration of the opportunistic approach to disinflation. *Journal of Monetary Economics* 53, 1877–1893.
- Andrews, B., Davis, R.A., Breidt, F.J. (2006). Maximum likelihood estimation for all-pass time series models. *Journal of Multivariate Analysis* 97, 1638–1659.
- Ball, L., and N.G. Mankiw (1994). Asymmetric price adjustment and economic fluctuations. *Economic Journal* 104, 247–261.
- Benati, L. (2008). Investigating inflation persistence across monetary regimes. *Quarterly Journal of Economics* 123, 1005–1060.
- Breidt, F.J., Davis, R.A., Lii, K.-S., Rosenblatt, M. (1991). Maximum likelihood estimation for noncausal autoregressive processes. *Journal of Multivariate Analysis* 36, 175–198.
- Cecchetti, S.G., and G. Debelle (2006). Has the inflation process changed? *Economic Policy*, April 2006, 311–352.
- Chib, S., and S. Ramamurthy (2011). DSGE Models with Student-t Errors. *Econometric Reviews* (forthcoming).
- Cúrdia, V., M. Del Negro, and D. L. Greenwald (2012). Rare Shocks, Great Recessions. *Federal Reserve Bank of New York Staff Reports*, No. 585.
- Fuhrer, J.C. (2010). Inflation persistence. In Friedman, B.M., and M. Woodford (eds.), *Handbook of Monetary Economics*, Volume 3. Elsevier.
- Gerlach, S. (2000). Asymmetric policy reactions and inflation. Unpublished mimeo. Bank for International Settlements.
- Hassler U., and B. Meller (2011). Detecting multiple breaks in long memory: The case of U.S. inflation. Discussion Paper Series 1: Economic Studies No. 26/2011, Deutsche Bundesbank.

- Karadi, P., and A. Reiff (2012). Large shocks in menu cost models. ECB Working Paper No. 1453.
- Koop, G., M.H. Pesaran, and S.M. Potter (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics* 74, 119–147.
- Kumar, M.S., and T. Okimoto (2007). Dynamics of persistence in international inflation rates. *Journal of Money, Credit and Banking* 39, 1457–1479.
- Lanne, M., A. Luoma, and J. Luoto (2012). Bayesian model selection and forecasting in noncausal autoregressive models. *Journal of Applied Econometrics* 27, 812–830.
- Lanne, M., and J. Luoto (2012). Has US inflation really become harder to forecast? *Economics Letters* 115, 383–386.
- Lanne, M., J. Luoto, and P. Saikkonen (2012). Optimal forecasting of noncausal autoregressive time series. *International Journal of Forecasting* 28, 623–631.
- Lanne, M., and P. Saikkonen (2011). Noncausal autoregressions for economic time series. *Journal of Time Series Econometrics* 3 (3), Article 2.
- Lanne, M., and P. Saikkonen (2013). Noncausal vector autoregression. *Econometric Theory* (forthcoming).
- Levin, A.T., and J.M. Piger (2003). Is inflation persistence intrinsic in industrial economies? ECB Working Paper No. 334.
- Lof, M. (2013). Noncausality and asset pricing. *Studies in Nonlinear Dynamics and Econometrics* (forthcoming).
- Nobay, B., I. Paya, and D.A. Peel (2010). Inflation dynamics in the U.S.: Global but not local mean reversion. *Journal of Money, Credit and Banking* 42, 135–150.
- Orphanides, A., and D.W. Wilcox (2002). The opportunistic approach to disinflation. *International Finance* 5, 47–71.

Pesaran, M.H., and S.M. Potter (1997). A floor and ceiling model of US output. *Journal of Economic Dynamics and Control* 21, 661–695.

Tillmann, P., and M.H. Wolters (2012). The changing dynamics of US inflation persistence: A quantile regression approach. Joint discussion paper series in economics, No. 05-2012. University of Marburg.

Tsong, C.-C., and C.-F. Lee (2011). Asymmetric inflation dynamics: Evidence from quantile regression analysis. *Journal of Macroeconomics* 33, 668–680.

van Dijk, D., P.H. Franses, and H.P. Boswijk (2007). Absorption of shock in nonlinear autoregressive models. *Computational Statistics & Data Analysis* 51, 4206–4226.

van Dijk, D., P.H. Franses, and A. Lucas (1999). Testing for ARCH in the presence of additive outliers. *Journal of Applied Econometrics* 14, 539–562.

Table 1: Estimation results of the third-order autoregressive models for the demeaned U.S. consumer price inflation.

	Model				
	AR(3,0)- $N$	AR(3,0)- $t$	AR(2,1)- $t$	AR(1,2)- $t$	AR(0,3)- $t$
$\phi_1$	0.338 (0.052)	0.302 (0.078)	0.419 (0.132)	0.949 (0.037)	
$\phi_2$	0.123 (0.055)	0.228 (0.084)	0.319 (0.102)		
$\phi_3$	0.330 (0.052)	0.298 (0.075)			
$\varphi_1$			0.031 (0.138)	-0.582 (0.087)	0.260 (0.098)
$\varphi_2$				-0.262 (0.073)	0.291 (0.069)
$\varphi_3$					0.224 (0.073)
$\sigma$		2.594 (0.413)	2.800 (0.480)	2.828 (0.622)	3.061 (1.072)
$\lambda$		3.452 (1.063)	3.294 (0.942)	2.978 (0.802)	2.597 (0.638)
Log-likelihood	-512.182	-366.032	-375.388	-368.211	-360.994
Ljung-Box (5)	0.395	0.182	<0.001	0.038	0.053
McLeod-Li (5)	0.969	0.926	0.541	0.001	0.127

AR( $r, s$ ) denotes the autoregressive model with the  $r$ th and  $s$ th order polynomials  $\phi(B)$  and  $\varphi(B^{-1})$ , respectively.  $N$  and  $t$  refer to Gaussian and  $t$ -distributed errors, respectively. The figures in parentheses are standard errors. Marginal significance levels of the Ljung-Box and McLeod-Li tests with 5 lags are reported.

Table 2: Absorption times of shocks of different size.

$\pi$	1970:1–2012:2			1970:1:1982:4			1983:1–2012:2					
	A	S	M	L	A	S	M	L	A	S	M	L
0.5	4.08	4.51	4.02	3.71	5.75	6.06	6.06	5.17	3.31	3.92	3.78	2.28
0.4	5.46	6.22	5.40	4.76	7.12	7.06	7.94	6.39	4.70	5.86	5.22	3.10
0.3	7.28	8.29	7.02	6.53	9.12	9.59	9.53	8.28	6.43	7.70	6.86	4.82
0.2	10.58	11.35	10.13	10.25	12.98	12.88	13.29	12.78	9.47	10.73	9.70	8.05
0.1	15.53	16.29	15.18	15.11	17.85	16.65	18.76	18.11	14.46	15.81	14.35	13.28

Each entry is the average  $\pi$ -life of all (A), small (S), medium-sized (M) and large (L) shocks (in absolute value) in each sample period. The cutoff points between the small and medium-sized shocks equal 0.78, 1.17, and 0.69, in the 1970:1–2012:3, 1970:1–1982:4, and 1983:1–2012:3 periods, respectively. The corresponding cutoff points between the medium-sized and large shocks equal 1.83, 2.53, and 1.56.

Table 3: Dependence of absorption times of shocks on initial inflation.

$\pi$	1970:1–2012:2			1970:1:1982:4			1983:1–2012:2		
	L	M	H	L	M	H	L	M	H
0.5	3.91	3.47	4.85	4.18	4.82	8.11	3.59	4.14	2.26
0.4	5.71	4.75	5.93	5.71	5.65	9.83	5.41	5.73	3.05
0.3	7.80	6.33	7.71	7.82	7.18	12.17	7.57	7.43	4.41
0.2	11.09	9.15	11.49	11.00	9.82	17.83	10.78	10.51	7.23
0.1	16.25	13.85	16.47	15.76	14.41	23.06	15.81	15.46	12.23

Each entry is the average  $\pi$ -life of a typical shock with low (L), medium (M) and high (H) inflation at the time of the shock in each sample period. The cutoff points between the low and medium inflation quarters equal 2.80, 4.97, and 2.25, in the 1970:1–2012:3, 1970:1–1982:4, and 1983:1–2012:3 periods, respectively. The corresponding cutoff points between the medium and high inflation quarters equal 4.59, 8.49, and 3.55.

Table 4: Average asymmetry measures for absorption times.

$\pi$	1970:1–2012:2		1970:1:1982:4		1983:1–2012:2	
	P	N	P	N	P	N
0.5	-0.87*	-0.47	-0.68	0.76	-0.98*	-0.17
0.4	-1.13*	-0.36	-0.32	-1.05	-1.60*	-0.03
0.3	-1.57*	-0.44	-0.32	-1.43	-2.90*	-0.10
0.2	-1.04*	-0.32	1.13	-1.29	-2.30*	-0.07
0.1	-1.24*	0.09	0.16	-1.33	-2.06*	-0.15

The entries are average  $ASYN(\pi, \delta_t, \omega_{t-1})$  measures over positive (P) and negative (N) shocks in each sample period. The numbers of positive (negative) shocks equal 84 (81), 31 (21), and 53 (60), in the 1970:1–2012:3, 1970:1–1982:4, and 1983:1–2012:3 periods, respectively. Entries larger than two times their standard error are marked with an asterisk.

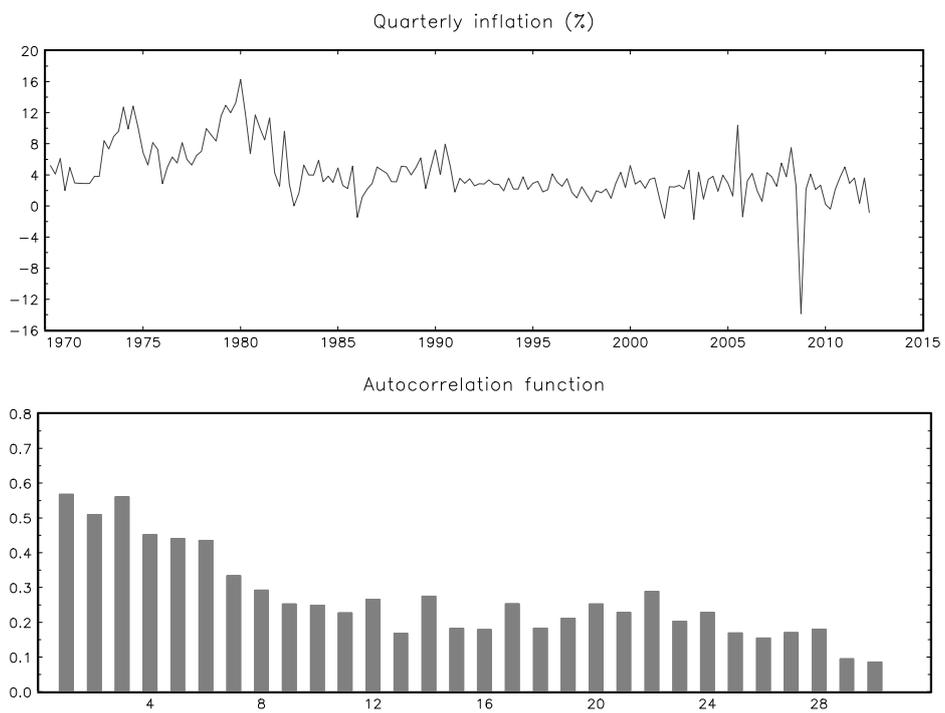


Figure 1: Quarterly U.S. consumer price inflation 1970:1–2012:2 and its autocorrelation function.

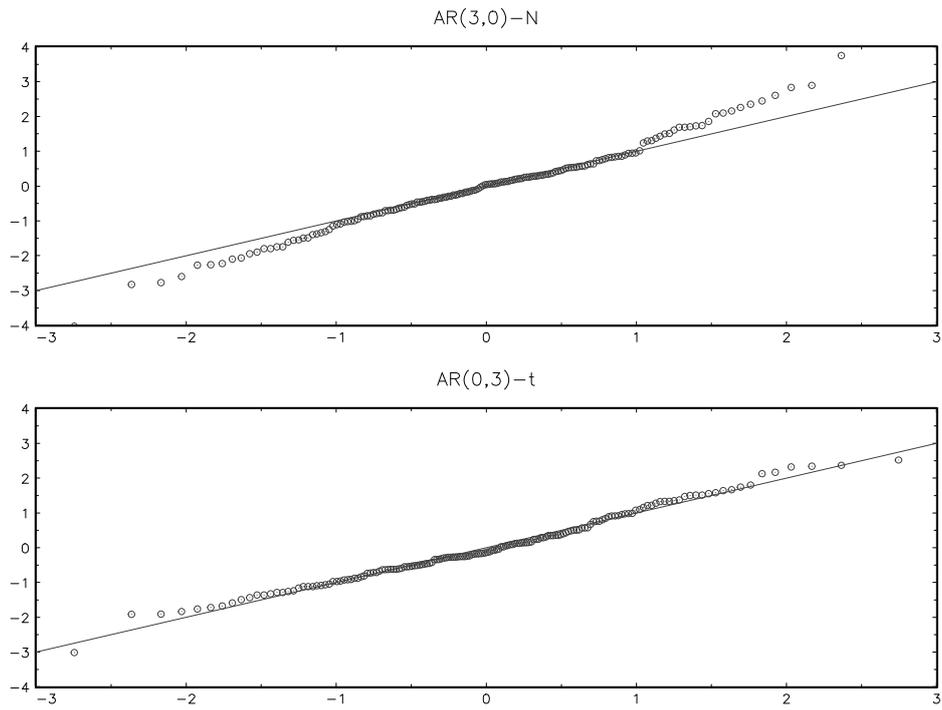


Figure 2: The quantile-quantile plots of the residuals of  $AR(3,0)-N$  and  $AR(0,3)-t$  models.

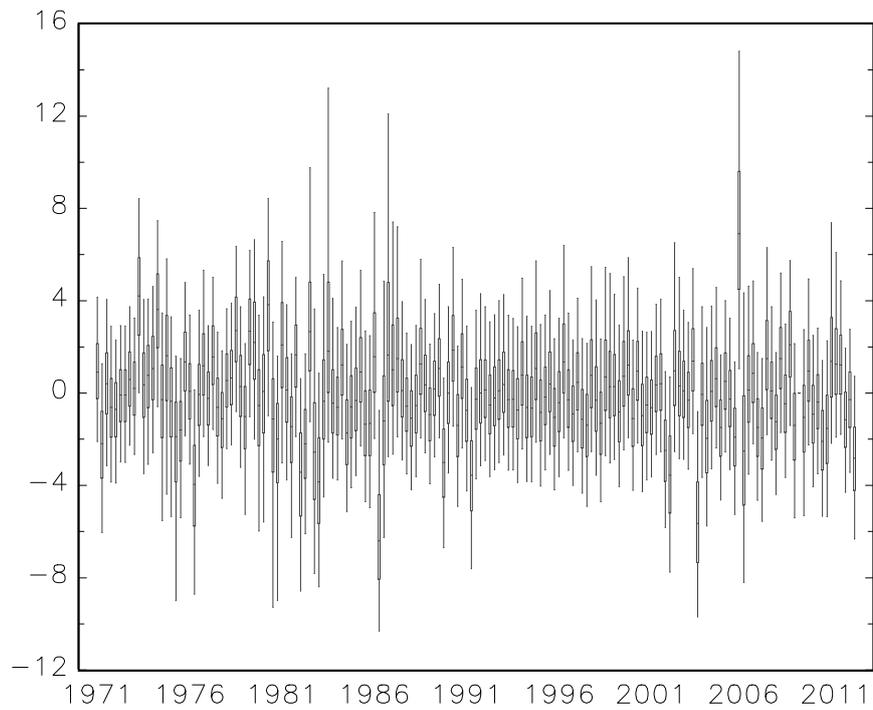


Figure 3: The box-and-whisker plot of one-step ahead forecast errors of the  $AR(0,3)-t$  model. The upper and lower bounds of each box are the 25th and 75th percentile of the predictive density, while the whiskers mark the 5th and 95th percentiles. The forecast error pertaining to 2008:4 is not included.